DOCUMENT RESUME

ED 087 795 TM 003 418

AUTHOR Thompson, Eugene W.
TITLE A Study of the Discrepancies Between Student Evaluation and Faculty Self-Perceptions of Instructional Procedures in Higher Education.
PUB DATE 72
NOTE 41p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Classroom Environment; *College Instruction; Course Evaluation; Effective Teaching; Evaluation Techniques; Higher Education; Role Perception; *Self Evaluation; *Student Opinion; *Teacher Evaluation; *Teaching Procedures; Teaching Skills
IDENTIFIERS *Student Opinions About Instructional Procedures

ABSTRACT

The study investigates the nature of the relationship between student evaluations and faculty self-perceptions of instructional procedures. Various characteristics of students and faculty were treated as independent variables in an effort to interpret the degree to which they affected the discrepancies between the two rating groups. The characteristics investigated were: student grade point average, class size, basis for course selection, and the amount of the instructor's teaching experience. The instrument used in the study was the Student Opinions About Instructional Procedures. The subjects for this study consisted of 58 instructors teaching 135 classes. Student and instructor responses to the questionnaire were used to develop discrepancy scores based on the three factors: professional competence, evaluation procedures, and student centeredness which the instrument measures. The data analyzed indicated that the independent variables tended to affect the outcomes in varying degrees. (Author)

ED 087795

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM 003 418
003 419

A STUDY OF THE DISCREPANCIES BETWEEN STUDENT EVALUATIONS AND FACULTY SELF-PERCEPTIONS OF INSTRUCTIONAL PROCEDURES IN HIGHER EDUCATION

Eugene W. Thompson
University of Alabama in Birmingham

Introduction

The evaluation of teaching performance has, in recent years, assumed a high level of priority in the minds of many of those concerned with the educational process, whether they be professionals or concerned laymen. In this project, interest is directed toward classroom behaviors and instructional procedures of higher educational instructors.

Of late, the procedure of utilizing student opinions has gained in favor to the point that many noted institutions now make systematic surveys of student opinions regarding instructional procedures. The source of the activity, however, has varied from place as has the preparation of a questionnaire. In some cases the entire procedure, including creation of the document, has been carried out by students who lacked measurement expertise, leading one to suspect the intent and reliability of the process. In others, the process originated and developed among faculty groups, generally departmental committees.

Assuming that development of such an instrument has taken place under the direction of persons possessing behavioral measurement skills, it would seem that there are two groups which could provide a reliable response. These groups or persons are those who would have first-hand knowledge of classroom interactions. First, of course, are the students and second the instructors. Others may have formed opinions regarding an instructor's classroom behaviors but these opinions will be based on rumor, hearsay, and only rarely on firsthand knowledge.

The purpose of this study was to examine the relationship between education instructors' self-perceptions of their instructional behaviors and procedures and their student's assessments of these same behaviors and procedures. The major interest was to determine then the congruent nature, if any, between the two sets of judgments.

An assumption was made that assessment of behaviors by individuals actually involved in the interaction represented one realistic approach to providing an evaluation tool which could provide feedback that could facilitate behavior change and teaching improvement. To meet this goal of a realistic model, 58 instructors in The Teacher Education Department, College of Education at Western Michigan University were requested to respond to an instrument designed to provide a rating of their self-perceived classroom behaviors and procedures. Similarly, over 3,000 students of participating instructors responded to a questionnaire which differed only by pronoun change in items of their perceptions of the instructor's behaviors and procedures. At the same time that these ratings were collected various personal characteristics of each respondent were also identified.

The examination of the degree of congruence between students and instructors was made by dichotomizing several variables of both student and instructor characteristics. These characteristics, it was hypothesized, would have an impact on the evaluation process. The resulting analyses serve as an indicator of the degree of congruence of student and self-evaluations as they compared on the basis of varying characteristics.

Definition of Terms

The use of the following terms in this study is intended to carry the attached definitions:

1. Questionnaire or instrument -- Data were collected from students by means of the Student Opinions About Instructional Procedures (SOAIP) a questionnaire explained more fully in a later section. Only minor pronoun modification was made to this instrument in order to adapt it to a means for collecting faculty data.
2. Student ratings (opinions) -- These ratings or opinions given by students were in regard to questions posed about instructor behaviors and instructional procedures as found in the SOAIP.
3. Instructor self-perceptions -- These self-ratings are also in response to the behaviors and procedures identified in the version of the instrument modified for faculty use.
4. Classroom behaviors and procedures -- The instructional behaviors and procedures referred to in this effort are those identified as relevant by faculty and students in the development of the questionnaire and are reflected in that instrument.
5. Achievement levels -- These levels are those reported in the student response to the questionnaire and are in the form of cumulative grade point averages. Higher achievement is referred to as category 3 (3.0 - 4.0), while lower achievement is considered to be in categories 0, 1, and 2 (under 1.0 through 2.99).
6. Class size -- Class size is stratified according to the reported number of student enrollment. This study will consider 1-30 students to be a small class, while one of 31 or over is large.

7. Student enrollment status -- This status for a course is reported by the student as being either elective or required.
8. Teaching experience -- These data are reported by the instructor, and refer to the number of years that faculty members have taught in higher education. A dichotomy has been established with "less experience" being category 0 (less than five years), and "more" category 1 (five years or more).

Review of Selected Related Literature

Eble (1970 p. 17) in a detailed report of a project designed to improve college teaching, lends considerable support to the uses of student evaluation processes for instructors. In support of this type of evaluation he cites the following arguments: (1) the changes are increased that excellence in teaching will be recognized, (2) greater students-teacher interaction may result, (3) the institution may consider its overall goal in light of this evaluation of teaching, (4) provision is thus made for the only direct information about faculty teaching, (5) a tangible sign is displayed by faculty of the need for student involvement in goal setting.

In an early study, Guthrie (1949) attempted to determine the correlation between student ratings and "faculty-jury" ratings of instructors using a nine item questionnaire dealing with general professional contributions but not specific classroom behaviors. His data indicated that while the correlation of student ratings with other student ratings were "of the order of .89", and between .64 and .76 when faculty-juries were compared with each other, the correlation between student and faculty-juries was .48. The statement was made that a likely cause of this

"radical difference" is the fact that students have sat through many hours with the instructor, while the faculty are highly dependent on personal acquaintance and student hearsay.

Maslow and Simmerman (1956) began an investigation out of what they referred to as "skepticism with the common tendency to dichotomize teaching and research (creating)." Correlating student and colleague ratings on instructors as good teacher, health personalities, and "creativeness in their field" the authors found a high degree of validity regarding student judgments when faculty judgments were used as a criterion measure. A high correlation ($r = .69$) with faculty judgments of the same teachers indicated to them that "a faculty cannot take student judgment lightly without casting aspersion on its "own competence to judge".

Checking claims for concurrent validity of student ratings Costin (1966) found $r = .49$ ($p. < .01$) between ratings assigned by students on an assessment of overal effectiveness. Low correlations were achieved, however, on the individual items of his scale. Reasoning that this result may be a function of the greater difficulty chairmen had in making judgments on individual items as compared to the "overall" category he concludes that the results do lend support to the claims of the validity of student ratings.

At the State College of Washington Downie (1952) worked with a sample of 300 students to investigate the relationship of student achievement, enrollment status, college rank, and the class size upon the assessment of four factors identified as: (1) instructional procedures, (2) exams, (3) cultural value, and (4) instructor-student relations. The data indicated that students with higher achievement rated instructors

higher on factor four. There was no difference between upper and lower division student ratings. Large classes, however, rated teachers lower on factor one, two, and three than did smaller classes. An additional independent variable investigated was instructor experience. No differences were found between those with experience above or below five year.

Working with 131 students at Clemson University, Caffrey (1969) found that course grades, overall GPA, and sex of students were not of critical importance in determining a student's rating of his teacher. Nor were the personal qualities of the teacher except for a factor labeled "rapport". The author decided that personal bias was absent from the evaluation of teachers made by students in this study.

Investigating numerous independent variables which could affect student ratings, McKeachie (1969) observed that students do not rate teachers on their personality but on how they are learning. Further findings indicated that undergraduate students tend to rate higher. Sex of the instructor had no effect on the ratings. Associate professors tended to be rated higher than other ranks. The degree held by the instructor did have an impact with higher degrees receiving higher ratings. Older teachers tended to be rated lower. The evidence was mixed regarding the effect of class size and whether the student was enrolled on a required or elective basis. The author made the statement that while student reactions are valuable for improving teaching he doubted their validity for inter-instructor comparisons. He further observed that the fact that feedback or knowledge of results aids learning is a "psychological principle of long standing".

With the rise in the use of student ratings of faculty and of the use of others' perceptions as feedback has come additional work on the effects

of feedback insofar as behavior change is concerned. As is often the case, the results of these studies have produced mixed evidence.

Tuckman and Oliver (1968) involved 286 teachers of vocational education at the high school and technical level. Some of the teachers received feedback from students, some from supervisors, some from both, and some none at all. Using a twelve week interval between pre and post testing, they concluded that student feedback had a significant effect but that from supervisors did not.

At Western Michigan University, Bryan (1963) spent approximately 35 years attempting to determine the effectiveness of written student feedback to teachers. His data indicated that 57 percent of the teachers receiving this feedback made significant gains in student ratings, compared to 24 percent of the control group.

A theory was developed by Gage, Runkle, and Chatterjee (1960) which they feel explains why behavior change is likely to occur when persons are provided with feedback from others. The rationale is that feedback can create an imbalance in the individual's self-perception that he will attempt to correct. His most likely response would be to modify his behavior, or at least to attempt to modify other's perceptions of his behavior.

Clark (1970) compared the effects of written student feedback, interaction analysis feedback, research-based statements, and group guidance in modifying the image of high school teachers. His work led him to conclude that all types of feedback were more effective in modifying teacher image than no feedback at all. Written student feedback was the most effective.

While empirical evidence concerning self-evaluation in education is almost non-existent, numerous authors have commented upon the need for, and value of, constant self-appraisal for teachers.

In his report of The Project To Improve College Teaching, Eble (1970, p. 11) commented that the peak of popularity has probably passed for the use of self-evaluation as a means of improving instruction. However, he observed that "self-evaluation which precisely sets forth a teacher's objectives might be achieved can be a brilliant contribution to our knowledge about teaching." He continued with the thought that this was probably not a sound idea for departmental evaluations.

In a study utilizing both rural and urban later-elementary school teachers, Amatora (1955) attempted to ascertain the relationships between self-ratings and ratings by peers for these teachers on a number of personality variables. Her highest correlations were found in traits such as persistence .54, thoughtfulness .47, common sense .45, and sense of humor .43. The lowest r's appeared on items such as intelligence .14, egotism .16, tolerance .18, and sincerity .21.

Studying the validity of self-estimate Shen (1925) compared self and peer ratings on eight traits. His work indicated a reliability range from .62 on impulsiveness to .91 for scholarship. He observed that the inaccuracy of self-estimates is largely due to systematic error on the part of the individual. This systematic error is believed to be a tendency to over or underestimate himself in all of the traits according to "the kind of delusions he has about himself".

In a study conducted in senior level educational psychology classes at the University of Delaware, Jenkins and Deno (1969) investigated the

effects of varying types of feedback upon student teacher self-evaluation. The treatment variations were either positive or negative feedback. Their data report led them to conclude that student classroom behavior has a powerful influence on self-evaluation by teachers, since significantly higher self-evaluation scores were recorded for those receiving positive feedback.

In the only study discovered which investigated the concept of comparing student ratings and self-evaluations, Webb and Nolan (1955) worked at the Naval Air Technical Training School at Jacksonville, Florida. Their study was based on feelings that "personal learning and improvement stems from an understanding of one's own adequacies and inadequacies". It was felt that a self-evaluation serves to focus the individual's attention on his inadequacies and as a result he will be motivated to attempt to correct them. They felt this must be carried out in a non-threatening situation and in conjunction with other evaluative procedures. Teachers in this study were largely not professional teachers. Likewise, supervisors were not trained professional supervisors. The data indicated a high relationship between student ratings and teacher self-ratings ($r = 62$). There was little relationship between either student or self-ratings with supervisor ratings. The authors stated that supervisors based their ratings on some factor other than those which were valid estimates of teaching ability or that they were random invalid intuition. "Not significant but high correlations (-.25 and -.23)" indicated that those teachers with higher measured intelligence and more education appeared to be more self-critical.

Populations and Samples

The population was located in the Teacher Education Department, College of Education, at Western Michigan University. The group was made up of all faculty members within the department, excluding those assigned to the Directed Teaching Office and the Continuing Education Center. Also, all students enrolled in Teacher Education Department (TEED) class offerings during Winter Semester, 1972 with the exceptions of the previously mentioned division, were part of the total population. These two divisions, Directed Teaching and Continuing Education, were eliminated because of the atypical nature of their functions within the department.

The 61 faculty members of the department who were teaching classes appropriate for the use of the Student Opinions About Instructional Procedures* questionnaire were approached in individual conferences and their cooperated solicited. Fifty-eight instructors agreed to participate and to supply their self-perceptions of their classroom behaviors. Thus, the faculty sample of the study was made up of 95 percent of the members of the population. These instructors taught a combined total of 135 classes. A characteristic of this sample to be noted, is that with the exception of two members new to the department, all participants had in the previous year received feedback data from students regarding their classroom procedures. Additional characteristics are presented in Table 1.

The student sample was determined by the cooperation of the faculty, since it was necessary to obtain each instructor's permission to utilize student data about himself. Thus, the student sample was made up of students enrolled in both graduate and undergraduate classes of participating

*Hereafter referred to as SOAIP.

instructors. The size of this sample was somewhat in excess of 3,000 students. A further description of this sample is presented in Table 2.

Instrumentation

Student Opinions About Instructional Procedures

The SOAIP questionnaire served as the main instrument; therefore, all data collected from students were derived from it. Developed in the Teacher Education Department at Western Michigan University through the cooperative efforts of faculty and students, the instrument has been used for three years as a means whereby students may evaluate teaching proficiency. Often such instruments are notable for the lack of systematic procedures by which they are produced. The SOAIP, however, was developed by an Ad Hoc Committee of the department with the specialty in education measurement being offered by Dr. Uldis Smidchens, Director of the Center for Educational Research at the University.

Work was begun on the questionnaire in September of 1969 following a charge presented by the University Faculty Senate to all departments for the development of instruments "to be used by students for the evaluation of teaching proficiency".

Using Senate guidelines, similar evaluation instruments used at Indiana University, Michigan State University, University of Michigan, and Ohio University were examined. Considerable overlap in the criteria used was noted. It appeared that the criteria found in these instruments would fit into one of the following categories: (1) evaluation of students, (2) personal relationships between faculty and students, (3) professional competence of faculty, and (4) individual (personal and physical) characteristics of the faculty member.

Under these categories, all statements from the instruments reviewed were then listed. These lists were reviewed and all redundant items deleted. From an original list of 140 criterion statements, a list of 61 relatively independent items were evolved. To involve faculty members and as many students and faculty selected from the 61 statements, both those which were meaningful to the evaluation of teaching proficiency, and those which could be evaluated. The thought behind this plan was that "any instrument which contained statements which were viewed by the faculty and students as being meaningful to the evaluation of teaching proficiency and further, were seen by the faculty and students as being measurable criteria, would be acceptable and valid for the population involved". (Committee for the Evaluation of Teaching Proficiency, 1970.)

These instruments of 61 items were then administered to both faculty and students. The faculty members were asked to make the judgments necessary and secondly, in terms of their interpretations of the value of the characteristics to the evaluation of teaching. Students were also requested to react to the items in two ways. First, to their own ability to rate faculty on the item characteristic and secondly, to their perception of the value of the item contents to the evaluation of teaching. A total of 61 faculty members were sent packages of material. Thirty-nine were returned representing a total of 923 usable student responses.

Responses of both faculty and students were analyzed to show percentages of both groups in each category on each item. The final instrument was developed by selecting only those items where there was at least 80 percent or greater agreement among both faculty and students on the importance of that item in the evaluation of teaching. The items selected were then rechecked in terms of the perceptions of faculty and

students' belief in the measurability of the item. Again, a minimum of 80 percent agreement was the cutoff point. Following this process, 21 items remained. Two were deleted because of redundancy. The remaining 19 items make up the evaluation form.

The instrument has been factor analyzed and three factors were identified which jointly account for 57 percent of the total variance. These factors have been labeled "Professional Competence, Evaluation Procedures, and Student Centeredness." In addition, a split-half reliability formula has been computed and has shown a Spearman-Brown r of .50 on the mean of the 19 items. While this value is somewhat low in view of previous studies reporting values in excess of .80 for similar instruments, it is felt from viewing the data that one reason for the lower value might lie with the scale used. The SOAIP utilizes a 5-point scale which describes the characteristic as almost never present, infrequently present, frequently present, almost always present, or undecided. Research conducted under direction of Roy C. Bryan at Western Michigan University by DuBois (1960 p. 28) indicates that a 4-point scale definitely tends to exhibit greater skewness than does a 5-point scale. The use of a 5-point scale produces a clear tendency toward a more normal frequency distribution. The use of the 5-point could have spread the scores out and possibly have produced the slightly lower reliability coefficient.

The Instructor Self-Perception instrument was used to record the instructors' self-perceptions about their instructional procedures. In the content of the individual 19 items, it is identical to the SOAIP. It differs only in that the pronouns have been changed from the third to first person. Informational items about the respondent vary also.

This study was designed to gather data from an existing field situation. Through analyzing the data by t-ratios and correlation coefficients it was believed that conclusions could be drawn regarding the relationships of the variables under consideration.

Near the end of the Winter 1972 semester, students responded to the SOAIP in terms of their class instructor. At approximately the same time, participating instructors recorded their own perceptions of their classroom procedures. The data from both of these groups were examined to determine the nature of the relationship between the dependent variables (rating scores) and independent variables (grade point average, class size, course selection basis, and instructor teaching experience). This examination of the independent-dependent variables was largely conducted through the use of discrepancy scores. These scores were developed simply by comparing each student's responses to the responses presented by the class instructor. These discrepancy scores represent the absolute difference between the faculty member and each of his/her students.

Statistical Analyses

All analyses undertaken were on the basis of the three factors the instrument was measuring. As means, discrepancy scores, t-ratios and correlations are discussed, it should be remembered that these scores and values were computed on the basis of each factor. For example, a mean score for each student from smaller classes was determined for factor one, factor two, and factor three, as were the means for all other categories and subcategories.

The data collected through the means described were analyzed by viewing the differences between various group means and through

Pearson Product-Moment Correlations. These findings are presented in table form in the following manner: First, characteristics of the samples included in the analyses are described and secondly, the results of the analyses are displayed. Statistically significant findings are denoted by an asterisk. Tables one and two describe the characteristics of the sample. Tables three, four, and five report the comparison of student and faculty response. Table six displays the correlations between factors as they were reported by students and faculty. Tables seven through eighteen make up the report of the data concerning the relationships between the independent variables and the discrepancies between student and faculty perceptions.

TABLE 1

Characteristics of Instructors Providing Self-Perceptions About Instructional Procedures

Total Number of Instructors	58	Rank of Instructor	
Total Number of Classes	135	Instructors	2
		Assistant Professors	22
Sex of Instructors		Associate Professors	18
		Professors	16
Male	38		
Female	20	Instructor Employment Status	
Degree Held by Instructor		Part-time	5
		Full-time	53
MA - MS	18		
Ed.S.	1	Instructor Teaching Experience	
Ed.D. - Ph.D.	39		
		Less than five years	16
		Five years or more	42

TABLE 2

Characteristics of Students Responding to the
Student Opinions About Instructional Procedures Questionnaire

Total Number of Students	
	3048
Sex of Students	
Male	997
Female	2004
No Sex Identified	47
Classification of Students	
Freshmen	48
Sophomores	721
Juniors	1055
Seniors	485
Graduates	638
Non-Degree	75
No Class Identified	26
Student Selection of Class	
Required	2419
Elective	588
No Selection Identified	41
Student Enrollment Status	
Part-time	592
Full-time	2424
Status Not Identified	32
Level Planning to Teach	
Elementary	1356
Secondary	924
Post-Secondary	193
Not Teaching But Education	222
Not Education	243
Level Not Identified	110
Reported Grade Point Average	
Less than 1.0	6
1.0 - 1.99	47
2.0 - 2.99	1260
3.0 - 4.0	1588
GPA Not Established	114
GPA Not Reported	33

TABLE 3

Relationship Between Student-Ratings and
Instructor Self-Perceptions Regarding Professional Competence

	Student Responses	Instructor Responses	df	t	p	r
N:	3024	135				
M:	2.656	2.663	3157	.172	.86	.04
SD:	.498	.314				

TABLE 4

Relationship Between Student Ratings and
Instructor Self-Perceptions Regarding Evaluation Procedures

	Student Responses	Instructor Responses	df	t	p	r
N:	2836	133				
M:	2.426	2.532	2967	1.726	.08	.05
SD:	.701	.591				

TABLE 5

Relationship Between Student Ratings and
Instructor Self-Perceptions Regarding Student Centeredness

	Student Responses	Instructor Responses	df	t	p	r
N:	3025	135				
M:	2.606	2.655	3158	1.011	.31	.12
SD:	.554	.277				

As no significant differences and little tendency to vary together between the variables were displayed, interest increased in the effort to learn more about the manner in which students were rating faculty and faculty were rating themselves. In Table 6 additional correlational values of interest are presented regarding these relationships. In it, the symbol S refers to students while I, indicates instructors. The letter F indicates factor. The numeral following this designation indicates the factor to which reference is being made according to the following code: (1) Professional Competence, (2) Evaluation Procedures, and (3) Student Centeredness.

TABLE 6

Pearson Product-Moment Correlations Between Student and Instructor Factor Ratings

	SF_1	SF_2	SF_3	IF_1	IF_2	IF_3
SF_1	1.000	.526	.785			
SF_2		1.000	.546			
SF_3			1.000			
IF_1				1.000	.234	.584
IF_2					1.000	.217
IF_3						1.000

Based on the findings reported in Table 6, it would appear that the student factor ratings show a fairly consistent tendency to vary together. Apparently, students tended to rate their instructors in a like manner concerning all classroom behaviors. In other words, an instructor rated high on professional competence is quite likely ($r = .78$) to be rated high on student centeredness, etc.

Instructors, on the other hand, while showing some of the same tendencies, appeared to discriminate more highly between the behavioral factors. Again, it would seem that instructors rated themselves highly ($r = .58$) on student centeredness and professional competence. As can be seen in Table 6, r was .58. Correlations, however, generally do not approach the strong relationships that were displayed by the student data.

TABLE 7

Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Professional Competence

	Discrepancy of High Achiever vs. Instructor	Discrepancy of Low Achiever vs. Instructor	df	t	p
N:	1576	1303			
M:	.371	.427	2777	3.602	.001*
SD:	.391	.436			

TABLE 8

Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perception Regarding Evaluation Procedures

	Discrepancy of High Achiever vs. Instructor	Discrepancy of Low Achiever vs. Instructor	df	t	p
N:	1459	1229			
M:	.612	.647	2686	1.420	.15
SD:	.637	.671			

TABLE 9

Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Student Centeredness

	Discrepancy of High Achiever vs. Instructor	Discrepancy of Low Achiever vs. Instructor	df	t	p
N:	1581	1298			
M:	.404	.432	2878	1.813	.07
SD:	.412	.438			

TABLE 10

Relationship Between Class Size and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Professional Competence

	Discrepancy of Small Class Students vs. Instructor	Discrepancy of Large Class Students vs. Instructor	df	t	p
N:	1339	1685			
M:	.426	.369	3022	3.767	.001*
SD:	.425	.403			

TABLE 11

Relationship Between Class Size and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Evaluation Procedures

	Discrepancy of Small Class Students vs. Instructor	Discrepancy of Large Class Students vs. Instructor	df	t	p
N:	1241	1575			
M:	.620	.631	2814	.417	.67
SD:	.651	.654			

TABLE 12

Relationship Between Class Size and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Student Centeredness

	Discrepancy of Small Class Students vs. Instructor	Discrepancy of Large Class Students vs. Instructor	df	t	p
N:	1340	1685			
M:	.450	.386	3023	4.120	.001*
SD:	.430	.412			

TABLE 13

Relationship Between Class Selection and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Professional Competence

	Discrepancy of Elective Class Students vs. Instructor	Discrepancy of Required Class Students vs. Instructor	df	t	p
N:	585	2401			
M:	.418	.386	2984	1.699	.08
SD:	.458	.458			

TABLE 14

Relationship Between Class Selection and the Congruence of Student Ratings with Faculty Self-Perception Regarding Evaluation Procedures

	Discrepancy of Elective Class Students vs. Instructor	Discrepancy of Required Class Students vs. Instructor	df	t	p
N:	532	2247			
M:	.586	.632	2777	1.462	.14
SD:	.602	.660			

TABLE 15

Relationship Between Class Selection and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Student Centeredness

	Discrepancy of Elective Class Students vs. Instructor	Discrepancy of Required Class Students vs. Instructor	df	t	p
N:	581	2405			
M:	.443	.407	2985	1.826	.07
SD:	.415	.423			

TABLE 16

Relationship Between Instructor Experience and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Professional Competence

	Discrepancy of High Experience Instructors vs. Students	Discrepancy of Low Experience Instructors vs. Students	df	t	p
N:	2180	815			
M:	.401	.371	2993	1.77	.08
SD:	.437	.351			

TABLE 17

Relationship Between Instructor Experience and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Evaluation Procedures

	Discrepancy of High Experience Instructors vs. Students	Discrepancy of Low Experience Instructors vs. Students	df	t	p
N:	2009	778			
M:	.670	.508	2785	5.898	.001*
SD:	.682	.563			

TABLE 18

Relationship Between Instructor Experience and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Student Centeredness

	Discrepancy of High Experience Instructors vs. Students	Discrepancy of Experience Instructors vs. Students	df	t	p
N:	2182	814	2994		
M:	.433	.363	2994	4.024	.001*
SD:	.428	.406			

Summary of the Data

As has often been the case in research dealing with student evaluations the findings present varying evidence. Not only have the various independent variables appeared to affect the discrepancy between student ratings and instructors' self-perceptions, but considerable differences were also noted according to the type of behavior being assessed.

Interestingly, no significant differences were noted in Tables 3, 4 and 5 between student ratings and instructor self-perceptions. However, while students and faculty were in close agreement regarding the assessment of Professional Competence and Student Centeredness behaviors, there was some discrepancy on Evaluation Procedures.

Student grade point average appeared to have little effect on the congruence between student ratings and the instructor's self-perceptions regarding his classroom behaviors. Only on the assessment of behaviors regarding Professional Competence was the variance between discrepancy scores at a statistically significant level.

The size of the class in which the instructional interaction took place appeared to have a highly significant impact on the congruence of student assessments and faculty self-perceptions regarding Professional Competence and Student Centeredness behaviors. In regard to both factors, the differences between the small and large classes were of such a size as to indicate change occurrence only one time in more than 1,000. No such effect was noted on the assessment of Evaluation Procedures.

No significant differences occurred in the discrepancy scores between students and instructors according to whether the student was enrolled on an elective or required basis. The differences noted in

discrepancy scores could have occurred by probability in the following order: Professional Competence .08, Evaluation Procedures .14, and Student Centeredness .07.

The independent variable, teaching experience related to instructors, also appeared to have considerable impact on the discrepancy levels between students and faculty as they perceived at least two types of behaviors. Highly statistically significant differences in the discrepancies between students versus instructors with high experience and students versus instructors with low experience regarding Evaluation Procedures and Student Centeredness behaviors were evident in the findings. The differences observed in the assessment of Professional Competence behaviors could have occurred with a probability of only eight in 100.

The evidence was inconsistent, but it does appear that certain characteristics of students and instructors alter the manner in which students rate and faculty self-perceive certain classroom behaviors. Viewing the results of the data analyses, a tendency is noted for the differences in discrepancies between student and faculty assessments to be greatest on the measurement of Professional Competence and Student Centeredness.

Student ratings of instructors in higher education have been reported to be both reliable and valid. In the opinion of this writer, these types of assessment combined with faculty self-ratings represent the most viable measures of the classroom interaction as far as the improvement of teaching is concerned.

Implications

Beyond the presentation of a little used evaluation tool this study also investigated additional variables in order to extend awareness

regarding the concept underlying the evaluation of teacher effectiveness. In order to achieve this goal, various characteristics of students and faculty were treated as independent variables and used to create dichotomies upon which the analyses were based. These findings extended knowledge of the various facets of the concept considerably by revealing that certain characteristics of both groups are areas to be treated with concern as this type of self-evaluation is conducted. For example, it was shown that the instructor's teaching experience could have a highly significant impact upon the degree of congruence between his/her and his/her student's ratings.

This tool is a very simple extension of the existing use of student feedback. The additional self-evaluation can provide a focus point for instructors upon discrepancies between their own and their students' perceptions. These discrepancies may or may not actually describe shortcomings in the instructor's behavior; that is for him/her to decide. The psychology literature, however, indicates that behavior is generally changed as the need for change is seen. The assessment process which was the basis of this study increased the opportunity for instructors to make the decision regarding the modification of classroom behaviors.

The SOAIP can be used by all instructors who are interested in determining what their students regard their effectiveness to be. Further, it provides a useful forum for students which lets them know that their opinions are sought and valued.

REFERENCES

Allen, D. The development of a clinic to improve university teaching. Unpublished Discussion Proposal Draft, University of Massachusetts, 1972.

Amatora, Sister Mary O. S. F. Validity in self-evaluation. Educational and Psychological Measurement, 1956, 16, 119-126.

Amatora, Sister Mary O. S. F. Self-appraisal in teaching personality. Journal of Educational Psychology, 1955, 46, 94-100.

Bendig, A. W. The use of student rating scales in the evaluation of instructors in introductory psychology. Journal of Educational Psychology, 1952, 43, 167-175.

Bendig, A. W. A preliminary study of the effect of academic level, sex, and course variables on student ratings of psychology instructors. Journal of Psychology, 1952, 34, 21-26.

Bendig, A. W. Relations of level of course achievement of students or instructor, and course ratings in introductory psychology. Educational and Psychological Measurement, 1953, 13, 437-448.

Borgatta, E. F. Student ratings of faculty. American Association of University Professors Bulletin, 1970, 56, 6-7.

Brown, J. W., Thornton, J. W. Jr. College teaching: A Systematic Approach. New York: McGraw-Hill, 1971.

Bryan, R. C. Reactions to Teachers by Students, Parents, and Administrators. Cooperative Research Project, No. 668, Sponsored by the United State Office of Education, Western Michigan University, Kalamazoo, Michigan, 1963.

Coffrey, B. Lack of bias in student evaluations of teachers. Proceedings of the 77th Annual Convention of the American Psychological Association, 1969, 4, 641-642.

Clark, F.A. The effects of selected forms of in-service training in modifying teacher image. Unpublished Doctoral Dissertation, Western Michigan University, 1970.

Cohen, A. M., Brawer, F. B. Measuring faculty performance. ERIC, Clearinghouse for Junior College Informance. American Association of Junior Colleges. Monograph Series, 1969.

Cohen, S. A., Berger, M. G. Dimensions of students rating of college instructors underlying subsequent achievement on course examinations. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 605-606.

Costin, F. Intercorrelations between students' and course chairmen's ratings of instructors. Unpublished Report, University of Illinois, 1966.

Crawford, P. L., Bradshaw, H. L. Perceptions of characteristics of effective university teachers: A scaling analysis. Educational and Psychological Measurement, 1968, 28, 1079-1085.

Daw, R. M., George, M. L. Effect of feedback from principals to teachers. Journal of Educational Psychology, 1967, 58, 131-138.

Downie, N. M. Student evaluation of faculty. Journal of Higher Education, 1952, 23, 495-496, 503.

DuBois, K. F. An investigation into the relative merits of two forms of a student reaction questionnaire. Unpublished Masters Thesis, Western Michigan University, 1960.

Eash, M. J., Bennett, C. M. The effect of class size on achievement and attitudes. American Educational Research Journal, 1964, 1, 229-239.

Eble, K. E. The recognition and evaluation of teaching. Project to Improve College Teaching. Washington, D.C.: American Association of University Professors and Association of American Colleges, 1970.

Eble, K. E. Professors as teachers. San Francisco: Jossey-Bass, Inc., 1972.

Festinger, L., Torrey, J., Willerman, B. Self-evaluation as a function of attraction to the group. Human Relations, 1954, 7, 161-174.

Furbay, A. Self-assessment groups as a means for improving teaching behavior. Contemporary Education, 1969, 40, 322-323.

Gadzella, B. M. College students views and ratings of an ideal professor. College and University, 1968, 44, 89-96.

Gage, N. L. The appraisal of college teaching. Journal of Higher Education, 1961, 32, 17-22.

Gage, N. L. Runkle, P. J., Chatterjee, B. B. Changing teacher behavior through feedback from pupils: An application of equilibrium theory. In W. W. Charters, Jr. and N. L. Gage (Ed.), Readings in the Social Psychology of Education. Boston: Allyn & Bacon, 1963.

Gage, N. L., Runkle, P. J., Chatterjee, B. B. Equilibrium theory and behavior change: an experiment in feedback from pupils to teachers. Urbana, Illinois: Bureau of Educational Research, Report No. 6 in the Studies in the Generality and Behavioral Correlates of Social Perception, 1960.

Gorwin, D. B., Payne, D. E. Evaluating instruction: Cross perceptions of college students and teachers. The School Review, 1962, 70, 207-219.

Gray, C. E. The teaching model and evaluation of teaching performance. Journal of Higher Education, 1969, 40, 636-642.

Guthrie, E. R. The evaluation of teaching. Educational Record, 1949, 30, 109-115.

Hayes, R. B. A way to measure classroom teaching effectiveness. Journal of Teacher Education, 1963, 14, 168-176.

Hudelson, E. The validity of student rating of instructors. School and Society, 1951, 73, 265-266.

Isaacson, R. L., McKeachie, W. J., Milholland, J. E. Correlation of teacher personality variables and student ratings. Journal of Educational Psychology, 1963, 54, 110-117.

Jarrett, J. L. The self-evaluating college teacher. Today's Education, 1969, 58, 40-41.

Jenkins, J. R., Deno, S. L. Influence of student behavior on teachers' self-evaluation. Journal of Educational Psychology. 1969, 60, 439-442.

Lauroesch, W. P., Pereira, P. D., Ryan, K. A. The use of student feedback in teacher training. Cooperative Research Project No. 8-E-115. Sponsored by the United States Office of Education, University of Chicago, Chicago, 1969.

Maslow, A. A., Zimmerman, W. College teaching ability, scholarly activity, and personality, Journal of Educational Psychology 1956, 47, 185-189.

McKeachie, W. J. Student ratings of faculty. American Association of University Professors Bulletin, 1969, 55, 439-444.

McKeachie, W. J., Lin, Y., Mann, W. Student ratings of teacher effectiveness: Validity studies. American Educational Research Journal, 1971, 8, 435-445.

Miller, M. T. Instructor attitude toward, and their use of student ratings of teachers. Journal of Educational Psychology, 1971, 62, 235-239.

Miller, R. I. Evaluating faculty performance. San Francisco: Jossey-Bass, Inc., 1972.

Powell, M. G. Comparisons of self-ratings, peer ratings, and experts' ratings of personality adjustment. Educational and Psychological Measurement, 1948, 8, 225-234.

Rayden, M. F. College student ratings of instructors. Journal of Experimental Education, 1968, 37, 76-81

Russell, H. E., Bendig, A. W. Investigation of the relations of student ratings of psychology instructors to their course achievement when academic aptitude is controlled. Educational and Psychological Measurement, 1953, 13, 626-635.

Ryan, K. A. The use of student written feedback in changing the behavior of beginning secondary school teachers. Unpublished Doctoral Dissertation, Stanford University, 1966.

Savage, M. L. Changes in student teachers through the use of pupil ratings. Unpublished Doctoral Dissertation, University of Illinois, 1957.

Shen, E. The validity of self-estimate. Journal of Educational Psychology, 1925, 16, 104-107.

Simpson, R. H. Teacher self-evaluation. New York: Macmillan Company, 1966.

Simpson, R. H. Instructors' use of self-evaluation procedures. Social Education, 1961, 25, 410-413.

Solomon, D. Teacher behavior dimensions, course characteristics, and student evaluation of teachers. American Educational Research Journal, 1966, 3, 35-47.

Stotland, E., Thorley, S., Thomas E., Cohen, A., Zander, A. The effects of group expectations and self-esteem upon self-evaluation. Journal of Abnormal and Social Psychology, 1968, 59, 297-301.

Voeks, V. W., French, G. M. Are student ratings affected by grades? Journal of Higher Education, 1960, 31, 330-334.

Weaver, C. H. Instructor rating by college students. Journal of Educational Psychology, 1960, 51, 21-25.

Webb, W. B., Nolan, C. Y. Student,supervisor, and self-ratings of instructional proficiency. Journal of Educational Psychology, 1955, 46, 42-46.

Wolthuis, R. M. A study of factors related to student ratings of college instructors. Unpublished Doctoral Dissertation, Western Michigan University, 1970.

APPENDIX A